April 24, 2023

Office of The Director
National Institutes of Health
9000 Rockville Pike
Bethesda, Maryland 20892

RE: ROR Comments in Response to NOT-OD-23-091, "Request for Information on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research"

To Whom It May Concern:

I write on behalf of the Research Organization Registry (ROR) responding to the Request for Information (RFI) on the NIH Plan to Enhance Public Access to the Results of NIH-Supported Research issued on February 21, 2023.

ROR is a global registry of open persistent identifiers for research organizations. ROR IDs are designed to be used in publication metadata and research infrastructure to unambiguously identify the organizations researchers are affiliated with, including their employers and their funders, so that people and systems can reliably connect research outputs to organizations. ROR is operated as a joint initiative by California Digital Library, Crossref, and DataCite, three not-for-profit organizations that have deep ties to research communities as well as extensive experience building and maintaining persistent identifier services and infrastructure.

ROR IDs are being integrated into various systems wherever there is a need to identify organizations and capture affiliation metadata. These implementations reflect the importance of affiliation metadata both upstream in the research and publishing process—i.e., identifying author affiliations upon submission of a manuscript—and in downstream services and systems for discovery and tracking of research, such as Crossref metadata, scholarly indexes and databases, and repositories. ROR is the preferred identifier for use in DOI metadata for publications, datasets, and grants registered in Crossref and DataCite, it is the primary identifier supported in ORCID records for researcher affiliations, and it has been recommended in national PID policies recently announced in Australia, Canada, the Netherlands, and the United Kingdom. ROR is unique among other organization identifiers because it is freely and openly available, specifically focused on connecting research organizations to research outputs, and designed to be used with other persistent identifiers, such as DOIs, ORCID IDs, and Funder Registry IDs.

ROR supports the NIH's interest in incorporating guidance on uses of PIDs and metadata in the Public Access Plan. Our specific comments on this aspect of the RFI notice are provided below:

> **4. Early input on considerations to increase findability and transparency of research.**
> *Section IV of the NIH Public Access Plan is a first step in developing the NIH's updated plan for persistent identifiers (PIDs) and metadata, which will be submitted to OSTP by December 31, 2024. NIH seeks suggestions on any specific issues that should be considered in efforts to improve use of PIDs and metadata, including information about experiences institutions and researchers have had with adoption of different identifiers.*

Persistent identifiers are an essential building block of research infrastructure. They facilitate disambiguation, enable discovery and tracking of research, and establish connections that can reveal key insights about how research is being conducted and consumed. While persistent identifiers on their own do provide a fundamentally useful function—unambiguous identification—they can be much more powerful and meaningful when they contain rich metadata and when they are linked to a network of multiple identifiers. This power is only unlocked when the identifiers and their underlying metadata are openly available for anyone to use and reuse, and it becomes especially relevant in computational contexts. ROR is a prime example of this power.

ROR IDs can be beneficial to NIH in many ways, including:

- Disambiguating and normalizing researcher affiliations
- Disambiguating and normalizing funder information
- Discovering and tracking research outputs associated with a specific institution
- Discovering and tracking research outputs associated with a specific funder or award
- Identifying connections between research awards, research funders, research outputs, and research organizations
- Enhancing the machine-readability and overall data quality of publication metadata, which supports accessibility needs as well as computational activities such as text and data mining of publications
- Facilitating the creation of automatic tools that track policy adherence by institutions

In order to realize these benefits, we encourage NIH to consider the following concrete actions:

- **Require or strongly encourage DOIs for NIH datasets and inclusion of ROR IDs in DOI metadata registered in DataCite**. ROR IDs can be included in DOI metadata for researcher affiliations, funder information, and publisher information (forthcoming in the next version of the DataCite metadata schema). This will enhance the discoverability of NIH data registered in DataCite, and make it possible for downstream discovery services to use this information to more efficiently track research outputs.

- **Register DOIs for NIH awards** (e.g., via Grant IDs provided by Crossref) and include ROR IDs in the award metadata so that downstream discovery services can use this information to more efficiently track research outputs connected to specific awards.

- **Require or strongly encourage investigators to obtain a DMP-ID for their Data Management and Sharing Plan**. This will ensure that the researcher's affiliation is automatically captured in the plan metadata in the form of a ROR ID and that the metadata about the plan, subsequent award, and resulting research outputs will be made publicly available in DataCite and downstream discovery services.

- **Map PubMedCentral author affiliations to ROR IDs and make this metadata available in PMC APIs**. This will make it possible to create more reliable search and browse features by author affiliation in PMC, to query PMC deposits to report on publications associated with specific institutions, and to connect the data to other indexes by normalizing on ROR IDs.

- **Encourage publishers to provide ROR IDs for author affiliations in DOI metadata registered with Crossref**. This will accelerate pressure on publishers to make their metadata openly available and will enable more efficient discovery and tracking of research outputs by institution.

Thank you for your consideration of these comments on behalf of ROR. We look forward to continued engagement on this issue.

Sincerely,

Maria Gould, ROR Lead
California Digital Library, University of California Office of the President